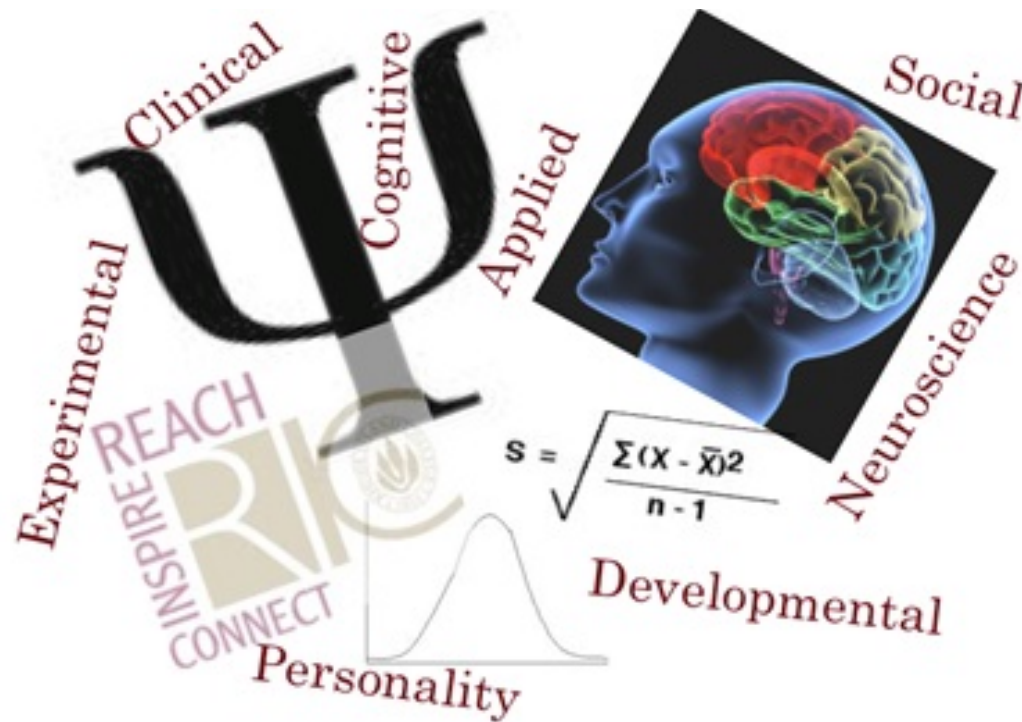


# PSYC202: Research Methods

Marcus Cappiello



# SCIENCE

- Simple idea: we should support claims with evidence!
  - Visual imagery mimics perception
    - Run and experiment
  - Waiting for difficult news is helped by support from friends
    - Observe people waiting
  - Cats are the cutest animal
    - Take a survey

# Theory, Hypotheses and predictions

- Theory
  - A systematic body of ideas about a topic or phenomenon
- Hypothesis
  - Based on a theory, come up with a *testable* idea or question

# Experiment

- Qualitative – observational, no statistics
  - Group discussions
  - Interviews
  - Small sample size
- Quantitative – experimental, run statistics
  - Need at least two conditions, one *experimental* condition and one *control* condition
  - Large sample size

# Experiment

- Between subject comparison
  - Two separate groups for the control and experimental conditions
    - Ex: Group A gets a new medicine, group B gets a placebo
- Within subject comparison
  - Same subjects participate in both conditions
    - Ex: Each participant takes a memory test underwater and in a classroom

# Experiment Tips

- Your experimental and control conditions should be as close as possible
  - Only difference is something that answers your question
  - Ex: ideal medical trial
    - Twins, give one drug and other placebo

# Experiment Tips

- Observer-expectancy effect
  - Researcher's bias influences the participants in the study
- Double blind
  - Both experimenter and subject don't know which condition a participant is in
  - Avoids experimenter bias

# Research article

- Abstract
- Introduction
- Method
- Results
- Discussion
- References
- Tables/figure captions/figures



# Treatment of Human Subjects History

- Nuremberg Code, 1949
  - First set of ethical codes for research on humans
- Declaration of Helsinki, 1964
  - 18<sup>th</sup> World Medical Assembly
  - Provided guidelines for the welfare of human subjects in medical research studies
- The Belmont Report, 1979

# Milgram's Obedience Experiment

- Yale University 1963, 1964, 1965
- “teacher” vs. “learner”
- “shock machine” from 15 to 450 volts



# Institutional Review Board

## Informed consent

- Based on Belmont Report--Three ethical principles
- Beneficence
  - Risks and benefits, 'Do no harm'
- Respect for persons (autonomy)
  - Informed consent
- Justice
  - Fair distribution of cost-benefits to subjects

# Variables

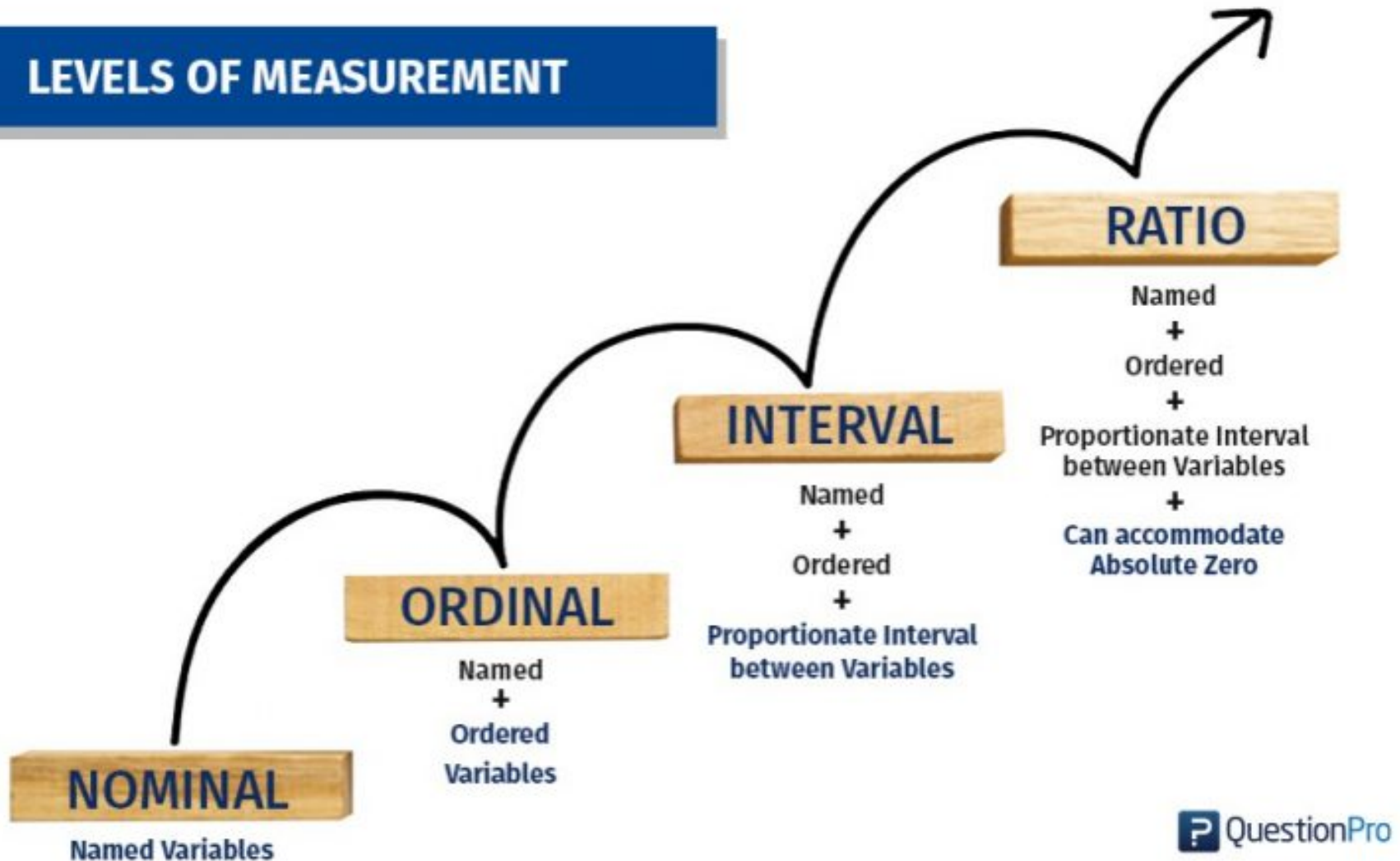
- Variable - Event, situation, behavior or characteristic which varies
- Dependent variable – measured output
  - Behavior/outcome due to independent variable
- Independent variable – what you manipulate or compare
  - Given drug or not (manipulate)
  - Left vs. right hand (compare)

# Operational definition

Specifying how variable(s) will be observed and/or measured in a study

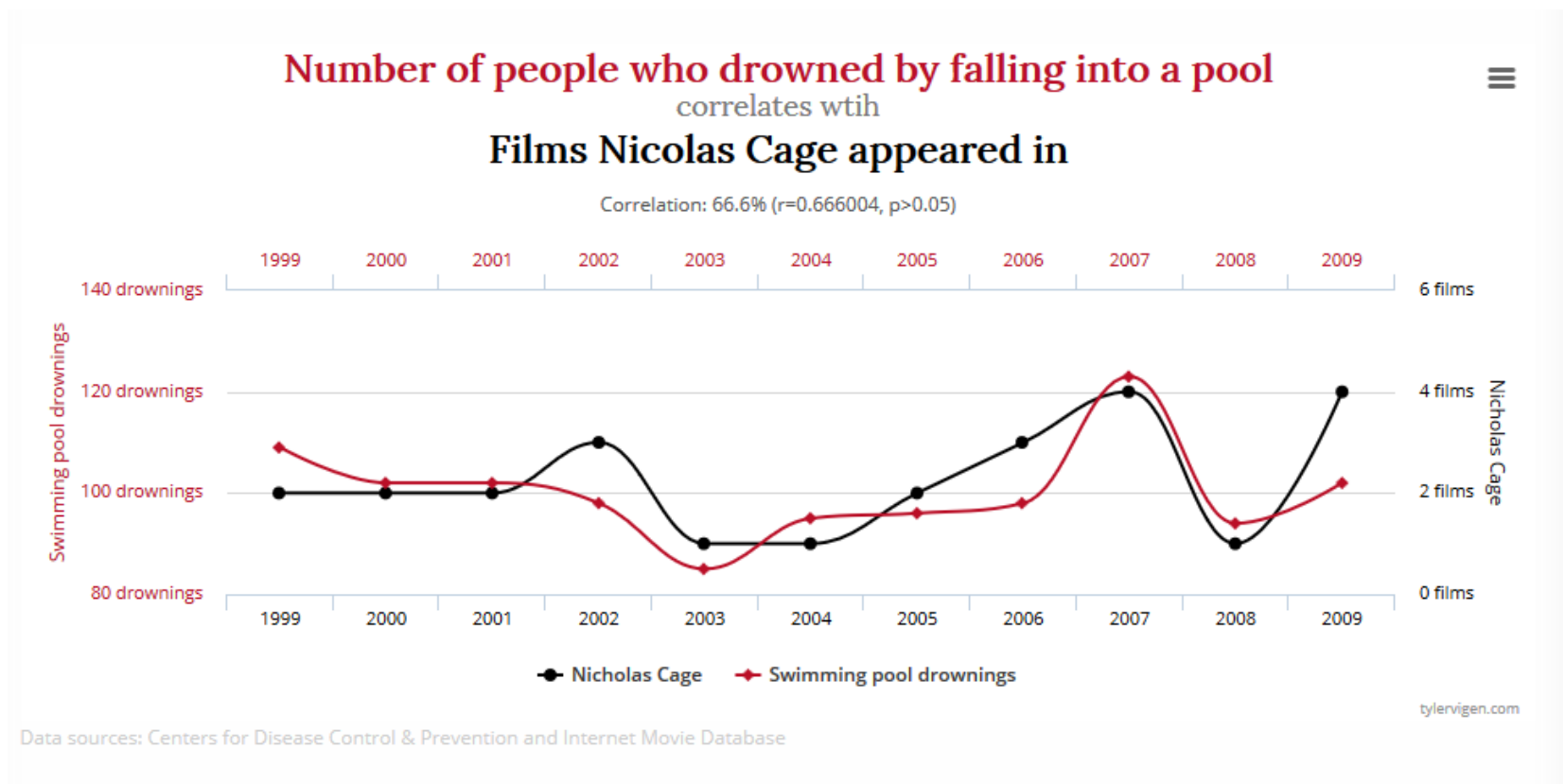
- Cognitive performance
  - Reaction time in ms to respond to stimulus
  - Accuracy
- Stress
  - Self report questionnaire, BP, cortisol
- Weight
  - Wt. in lb. using a spring scale with participants fully undressed after 10 hrs. of fasting

# Types of Variables



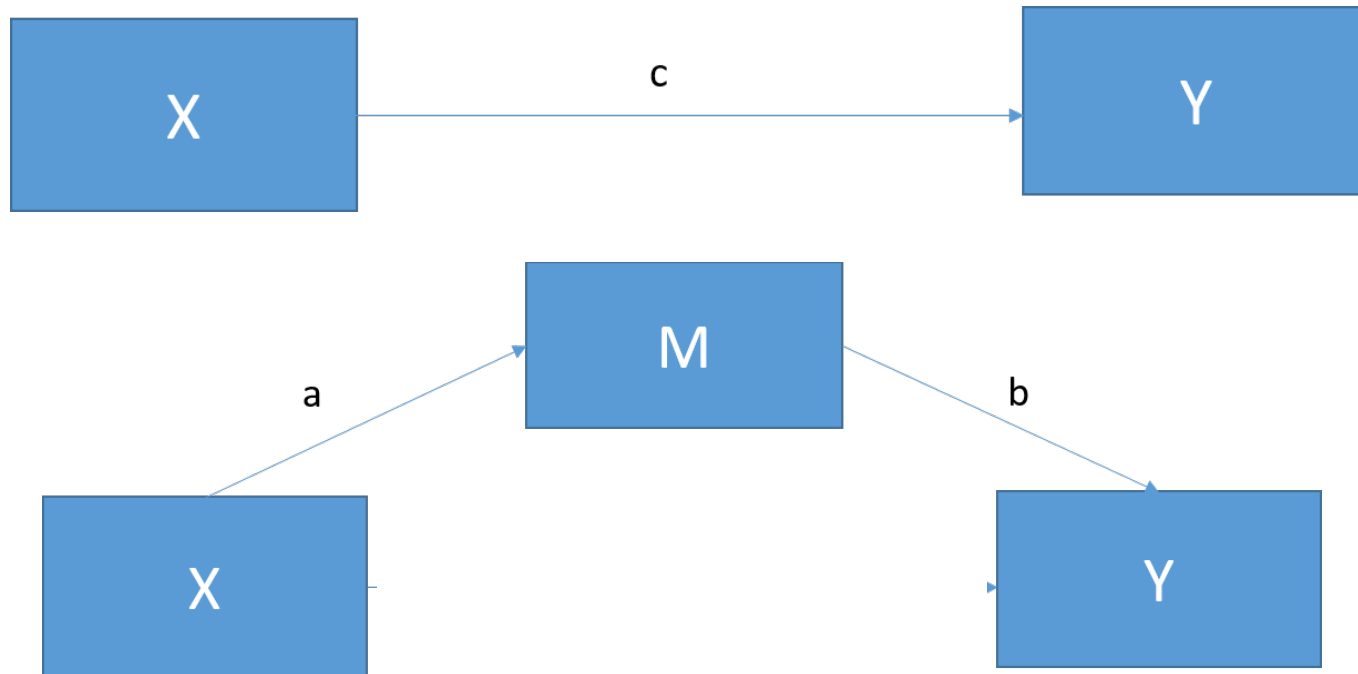
# Cognitive neuroscience methods

## ➤ Correlation versus causation



Do Nicolas cage movies *cause* swimming pool drownings?

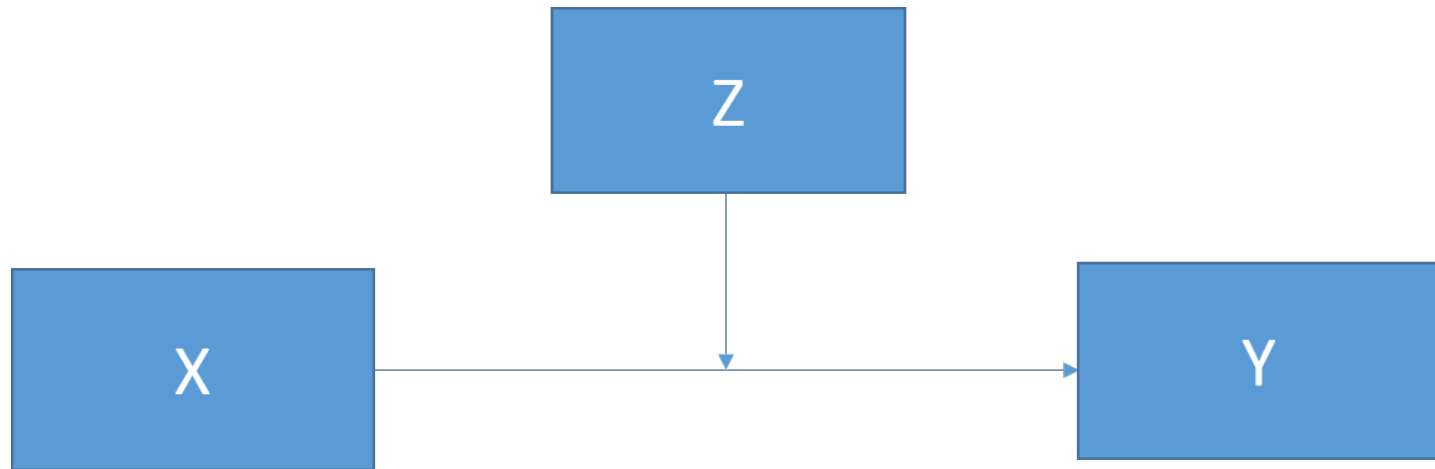
# Mediation



- Mediation – how/why a relationship exists.
  - Mediator is the causal result of X (IV) and causal antecedent of Y (DV)
  - Example:
    - X – grades
    - Y – happiness
    - M – self-esteem



# Moderation



- Moderation – affects the relationship itself
  - NOT the causal result of X
  - Example:
    - X – amount of time studying
    - Y – grades
    - Z – grade level (elementary or college)

# Quantitative vs. Qualitative Research

- Naturalistic observation
  - ✓ Participation and concealment
  - ✓ Limits
    - Difficult
    - Lack of control
- <https://www.youtube.com/watch?v=CWnk6PTsZNo>

# Other types of research

- o Case Study
  - v Phineas Gage
  - v K.C.
  - v 10-second Tom
  - v Split-brain patients
  - v Oliver Sacks

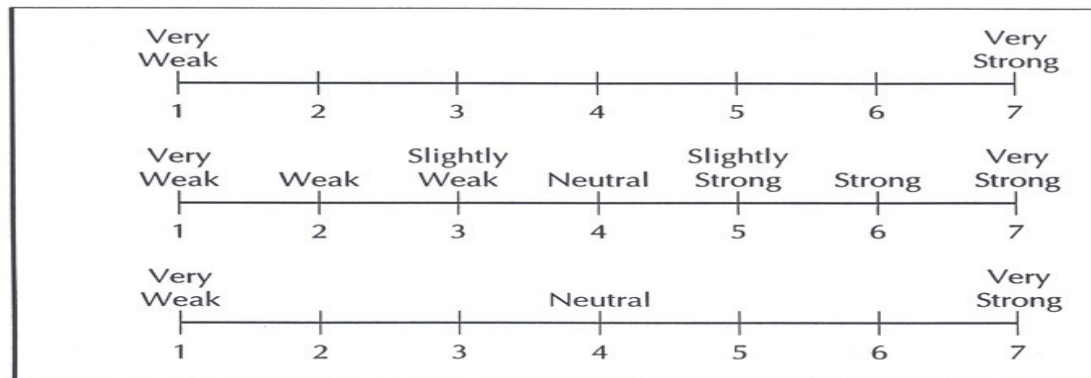


# Question wording

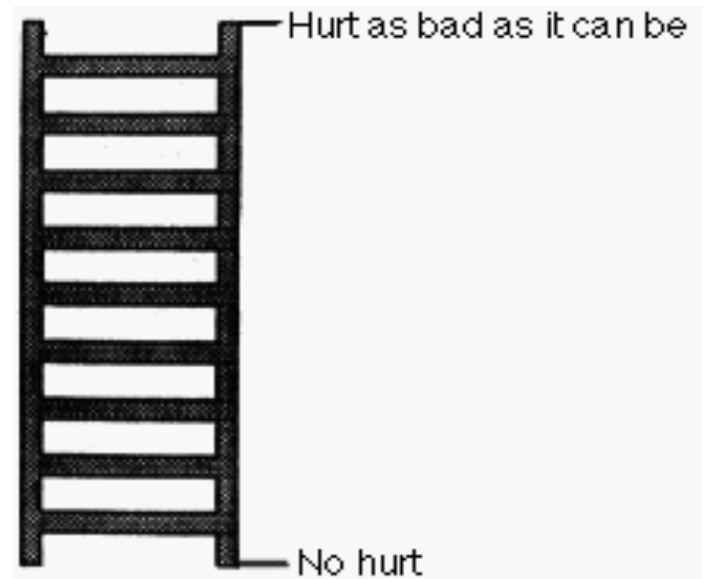
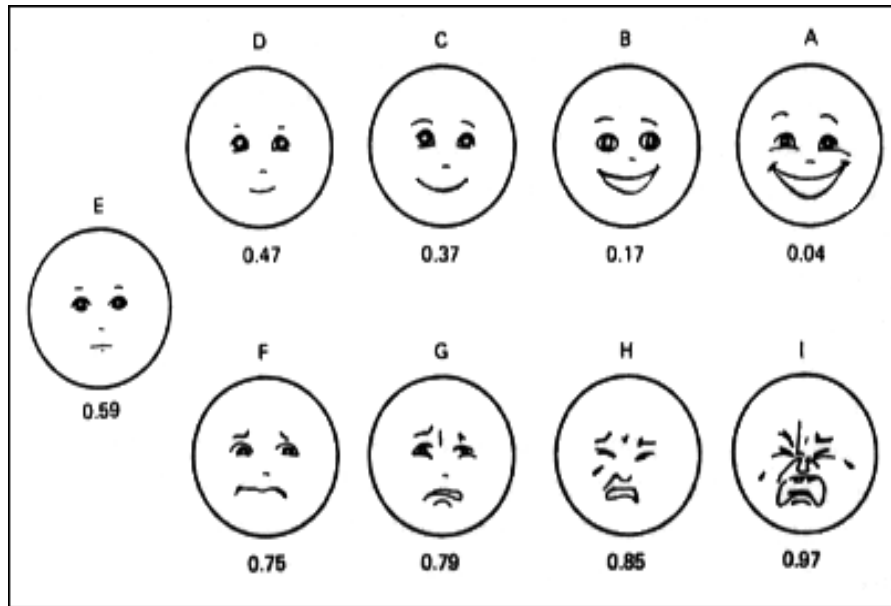
- Simplicity/clarity
- Double-barreled questions
- Loaded questions
- Negative wording
- Yea- vs. nay-saying

# Responses to questions

- Closed vs. open-ended responses
  - ✓ Quantitative vs. qualitative research
- Closed-ended responses
  - ✓ Yes/no or nominal vs. scale
  - ✓ Likert/rating scale
    - Restriction of range (1 to 3 vs. 1 to 10 response)



# Nonverbal Scales



# Interviews

- Face-to-face interviews
  - ✓ Interview bias
- Telephone interviews
- Focus group interviews
- Panel studies (interviews at multiple points of time)

# Probability sampling

- Simple random sampling
  - ✓ Every member of population has equal opportunity of being selected for the sample
- Stratified random
  - ✓ Control for particular variables to ensure they are equal across groups
- Cluster sampling
  - ✓ Identify clusters and then randomly sample from cluster



# Nonprobability sampling

- Haphazard sampling
  - ✓ “convenience”
  - ✓ Take them where you find them
- Purposive sampling
  - ✓ Sample meets certain criterion
- Quota sampling
  - ✓ Certain percentage of subgroups

# Convenience Samples

## aka Nonprobability sampling

---

- Assess representativeness of sample
- Cost and time trade-offs
- Relationships between variables vs. estimating population values

# Reliability

- The degree to which our measurements are consistent and contain a minimum of measurement error
  - If you weigh yourself multiple times during the day, do you get the same result?

# Types of Reliability

## INTERNAL

(extent to which a measure is consistent within itself.)

### **split-half method:**

measures the extent to which all parts of the test contribute equally to what is being measured.

## EXTERNAL

(the extent to which a measure varies from one use to another.)

**test re-test:** measures the stability of a test over time.

**Inter-rater:** to the degree to which different raters give consistent estimates of the same behavior

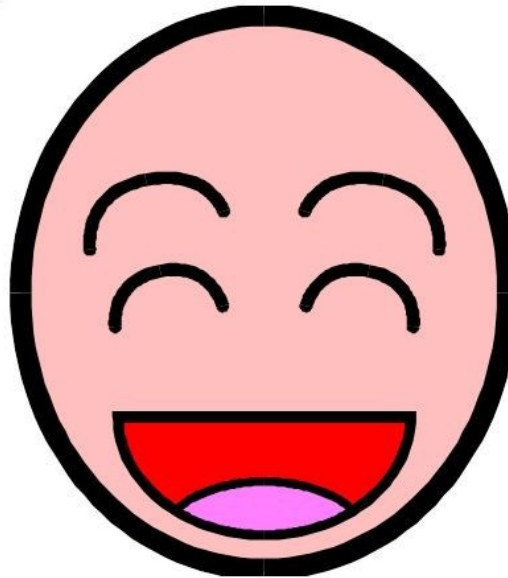


# Validity

- The extent to which a procedure/instrument measures what it is intended to measure.

# Face Validity

## Face Validity



Looks Good To Me

- It looks OK
- Looks to measure what it is supposed to measure
- Look at items for appropriateness
  - Client
  - Sample respondents
- Least scientific validity measure

# Content Validity

Does measure assess appropriate content?

## Content validity

Spider phobia

### Aspects of the construct

Strength of fear reaction

Persistence of reaction

Invariability of reaction

Recognition that reaction is unreasonable

Avoidance of spiders

...

### Aspects assessed



# Criterion Validity (Crit)

Do the scores correlate with other variables you would expect them to correlate to?

Example: test anxiety

- Should negatively correlate with test performance
- If not, it may be due to your questionnaire, rather than a novel finding





# Convergent vs. Discriminant

## Convergent and Discriminant Validity

- To have construct validity, a measure should both:
- Correlate with other measures that it should be related to (**convergent validity**)
- And, not correlate with measures that it should not correlate with (**discriminant validity**)

# 2X2 Design – Simplest factorial

	Independent variable A: Confederate sociability	
Independent variable B: Confederate food intake	Sociable	Unsociable
Low	Sociable/ low food intake	Unsociable/ low food intake
High	Sociable/ high food intake	Unsociable/ high food intake

# Percent errors in recall

## Main effects

Questioner type	Type of question		Overall means (questioner)
	Unbiased	Misleading	
Knowledge	13	13	
Naïve	13	13	
Overall means (Q type)			

# Percent errors in recall

## Main effects

Questioner type	Type of question		Overall means (questioner)
	Unbiased	Misleading	
Knowledge	20	13	
Naïve	13	13	
Overall means (Q type)			

# Percent errors in recall

## Main effects

Questioner type	Type of question		Overall means (questioner)
	Unbiased	Misleading	
Knowledge	20	13	
Naïve	13	20	
Overall means (Q type)			

# *F*-tests (ANOVA)

- More than two groups or more than one independent variable:
  - Analysis of Variance (ANOVA)
  - $F(d,f)$  = test statistic,  $p < .05$

# *F*-tests (con' t.)

- One IV, more than two levels
  - One-way ANOVA, (remember post-hoc test)
- Two or more IV' s, between-subject
  - Univariate ANOVA (Factorial ANOVA)
- Two or more IV' s, at least one IV within-subject
  - Repeated measures ANOVA
  - Ex: measure effect at three different time points

# Errors in Decision

## Type I and Type II

- Type I error
  - We reject the null hypothesis when it is actually true.
  - When  $\alpha = .05$ , chance of making a type I error is 5%



# Errors in Decision (con't.)

## Type I and Type II

- Type II error
  - Population means are truly different but results of our experiment do not lead to decision to reject the null
- Probability of type II error is beta or  $\beta$

# Power

- The probability of rejecting a false null hypothesis ( $H_0$ ).
- The probability of obtaining a value of  $t$  that is large enough to reject  $H_0$  when  $H_0$  is actually false
- We always test the null hypothesis against an alternative/research hypothesis

# Type II error (con't.)

- Probability of type II error related to
  - Alpha level
  - Sample size
  - Effect size—the magnitude of the effect
- Power =  $1 - \beta$ 
  - or  $1 -$  probability of type II error

# Factors Affecting Power

1. Size of the effect    ↑
2. Standard deviation of the characteristic    ↓
3. Bigger sample size    ↑
4. Significance level desired    ↓

# In practice

- You need:
  - Alpha level (0.05)
  - Sought power (usually 80%)
  - Standard deviation
    - Estimate based on a pilot study or past literature
  - Effect size!
- <https://clincalc.com/stats/samplesize.aspx>

# Effect size

- Difference between means
- For power, what is the smallest effect size of scientific interest?
  - Ex: Is a 1% decrease in anxiety meaningful? 10%?
  - Ex: Is a 40ms change in reaction time meaningful? 1ms?
- Subjective so be conservative

# Statistical issues

## Typical standards

- Alpha levels of .05, two-tailed
  - Willingness to accept that something is significant when it's really not 5% of the time
- Bonferroni corrections for multiple statistical tests
  - If 10 tests, alpha level for each test should be .005
- Power of .80 or greater

# Introduction section

- First paragraph – a hook and briefly what you will be talking about
- Middle paragraphs – discuss each aspect of your hypothesis, then your justification
- Final paragraph(s) –
  - Re-state your justification in a few sentences
  - State your hypothesis
  - Brief overview of experiment if writing a proposal



# The Logic of Meta-analysis

- Traditional methods of review focus on statistical significance testing
- Significance testing is not well suited to this task
  - Highly dependent on sample size
  - Null finding does not carry the same “weight” as a significant finding
    - significant effect is a strong conclusion
    - nonsignificant effect is a weak conclusion
- File drawer issue
- Meta-analysis focuses on the **direction** and **magnitude** of the effects across studies, not statistical significance
  - Isn't this what we are interested in anyway?
  - Direction and magnitude are represented by the effect size

# Effect Size: The Key to Meta-analysis

- The effect size makes meta-analysis possible
  - It is the “dependent variable”
  - It standardizes findings across studies such that they can be directly compared

# Which Studies to Include?

- It is critical to have an explicit inclusion and exclusion criteria
  - The broader the research domain, the more detailed they tend to become
  - Refine criteria as you interact with the literature
  - Components of a detailed criteria
    - distinguishing features
    - research respondents
    - key variables
    - research methods
    - cultural and linguistic range
    - time frame
    - publication types

# Strengths of Meta-analysis

- Imposes a discipline on the process of summing up research findings
- Represents findings in a more differentiated and sophisticated manner than conventional reviews
- Capable of finding relationships across studies that are obscured in other approaches
- Protects against over-interpreting differences across studies
- Can handle a large numbers of studies (this would overwhelm traditional approaches to review)

# Weaknesses of Meta-analysis

- Requires a good deal of effort
- Mechanical aspects don't lend themselves to capturing more qualitative distinctions between studies
- Selection bias poses a continual threat
  - ❑ Negative and null finding studies that you were unable to find
  - ❑ Outcomes for which there were negative or null findings that were not reported
- Analysis of between study differences is fundamentally correlational

## b). Effect 2: ADRA2B Effect in the Negative (vs. Neutral) Condition

